CLAIMS

What is claimed is:

1.   A system for distributed control of a communication switch, including a switch fabric and a plurality of switch satellites, in which the switch satellites include interfaces to external communication channels and interfaces to the switch fabric, and maintain one or more ingress queues for ingress communications from the external communication channels to the switch fabric, and one or more egress queues for egress communications from the switch fabric to the external communication channels, comprising:

a plurality of arbitration logic units coupled with respective switch satellites, a particular arbitration logic unit in the plurality of arbitration logic units including logic to control an arbitration cycle for a given switch cycle, including

a first stage in which performance parameters are gathered from other arbitration logic units in the plurality, the performance parameters indicating a status of the one or more egress queues maintained in the switch satellites coupled with the respective arbitration logic units,

a second stage in which bid data are propagated among the plurality, the bid data including a set of bids for use of egress queues during the switch cycle, the bids in the set including a destination identifier indicating a destination egress queue in one of the plurality of switch satellites, and a pressure parameter indicating a result of a combination of the performance parameter of the destination egress queue with a condition of a source ingress queue,

a third stage in which a connection map based on the bidding is computed, and

a fourth stage in which the switch fabric is configured, based upon the connection map.

2.   The system of claim 1, wherein said bids include priority parameters, and wherein said pressure parameter is a function of said priority parameters.

3.   The system of claim 1, wherein the plurality of arbitration logic units include configuration logic indicating a bid order, and the second stage includes a first arbitration logic unit in the order sending bid data to a next arbitration logic unit in the order, the next arbitration unit in the order consolidating and sending bid data to a next, and so on until the last arbitration logic unit in the order receives consolidated bid data, the bid data in the respective arbitration logic units being based upon the gathered performance data and conditions of ingress queues maintained in the respective arbitration logic unit and any previous arbitration logic unit, or units, in the order.

4.   The system of claim 3, wherein the third stage includes completing the formation of the connection map at the last arbitration logic unit in the order.

5.   The system of claim 4, wherein the third stage includes broadcasting the connection map from the last arbitration logic unit in the order to the other arbitration logic units in the plurality.

6.   The system of claim 1, wherein the switch fabric comprises a crossbar switch.

7.   The system of claim 1, including control communication logic coupled with the switch fabric supporting communication among the plurality of arbitration logic units.

8.   The system of claim 1, including control communication logic coupled with the switch fabric synchronizing communication among the plurality of arbitration logic units during an arbitration cycle.

9.   The system of claim 1, including control communication logic coupled with the switch fabric supporting communication of the connection map to the switch fabric from at least one of the plurality of arbitration logic units.

10.   The system of claim 1, wherein the performance parameter comprises a measure of backpressure.

- 16 -

11.   The system of claim 1, wherein said plurality of arbitration logic units support a source overload condition, and source biasing is reduce by randomizing the order of the set of bids.

12.   A method for distributed control of a communication switch, including a switch fabric and a plurality of switch satellites, in which the switch satellites include interfaces to external communication channels and interfaces to the switch fabric, and maintain one or more ingress queues for ingress communications from the external communication channels to the switch fabric, and one or more egress queues for egress communications from the switch fabric to the external communication channels, comprising:

gathering performance parameters in each switch satellite in the plurality from other switch satellites in the plurality, the performance parameters indicating a status of the one or more egress queues maintained in the respective switch satellites, sharing bid data among switch satellites in the plurality, the bid data including a set of bids for use of egress queues during the switch cycle, the bids in the set including a destination identifier indicating a destination egress queue in one of the plurality of switch satellites, and a pressure parameter indicating a result of a combination of the performance parameter of the destination egress queue with a condition of a source ingress queue, computing a connection map based on the bidding in at least one of the switch satellites in the plurality, and configuring the switch fabric based upon the connection map.

13.   The method of claim 12, wherein said bids include priority parameters, and wherein said pressure parameter is a function of said priority parameters.

14.   The method of claim 12, including indicating a bid order of the switch satellites, and sharing the bid data includes sending bid data to a next switch satellite in the order, the next switch satellite in the order consolidating and sending bid data to a next, and so on until the last switch satellite in the order receives consolidated bid data, the bid data in the respective switch satellites being based upon the gathered performance data and conditions of ingress queues

- 17 -

6   maintained in the respective switch satellite and any previous switch satellite, or satellites, in the

7   order.

1   15.   The method of claim 14, including completing the formation of the connection

2   map at the last switch satellite in the order.

1   16.   The method of claim 15, including broadcasting the connection map from the last

2   switch satellite in the order to the other switch satellites in the plurality.

1   17.   The method of claim 12, wherein the switch fabric comprises a crossbar switch.

1   18.   The method of claim 12, including supporting communication among the

2   plurality of switch satellites through the switch fabric.

1   19.   The method of claim 12, including synchronizing communication among the

2   plurality of switch satellites during an arbitration cycle.

1   20.   The method of claim 12, including supporting communication of the connection

2   map to the switch fabric from at least one of the plurality of switch satellites.

1   21.   The method of claim 12, wherein the performance parameter comprises a

2   measure of backpressure.

1   22.   The method of claim 12, including supporting a source overload condition, and

2   reducing source bias by randomizing the order of the sets of bids.

1   23.   A method for distributed control of a communication switch, including a crossbar

2   switch and a plurality of switch satellites, in which the switch satellites include interfaces to

3   external communication channels and interfaces to the switch fabric, and maintain one or more

4   ingress queues for ingress communications from the external communication channels to the

5   crossbar switch, and one or more egress queues for egress communications from the crossbar

6 switch to the external communication channels, where an ingress queue constitutes virtual

7 output queue for one of the one or more egress queues, the method comprising:

8 gathering performance parameters in each switch satellite in the plurality from other

9 switch satellites in the plurality, the performance parameters indicating a

10 backpressure of the one or more egress queues maintained in the respective

11 switch satellites;

12 sharing a bid data matrix among switch satellites in the plurality, the bid data matrix

13 including a set of bids including one bid for each of the virtual output queues for

14 use of egress queues during the switch cycle, the bids in the set including a

15 weighted pressure parameter indicating a result of a combination of the

16 performance parameter of the destination egress queue with a condition of the

17 ingress queue of the virtual output queue;

18 computing a connection map based on the bidding in at least one of the switch satellites

19 in the plurality; and

20 transmitting a vector to the crossbar switch for configuring the crossbar switch based

21 upon the connection map.

1 24.   The method of claim 23, wherein the bid data matrix comprises a data structure

2 for holding a bid for each of the virtual output queues serviced by the crossbar switch, and

3 including indicating a bid order of the switch satellites, and sharing the bid data matrix includes

4 computing bid data matrix in a first switch satellite having entries for virtual output queues

5 which originate in the first switch satellite, sending the bid data matrix to a next switch satellite

6 in the order, the next switch satellite in the order re-computing the bid data matrix with entries

7 for virtual output queues originating in the said first switch satellite and said next switch

8 satellite, and sending bid data to a next, and so on until the last switch satellite in the order

9 receives consolidated bid data and finally computes the bid data matrix with entries for virtual

10 output queues originating in all the switch satellites.

1 25.   The method of claim 23, including supporting communication among the

2 plurality of switch satellites through the switch fabric.

- 19 -

1   26.   The method of claim 23, including synchronizing communication among the
2   plurality of switch satellites during an arbitration cycle.

1   27.   The method of claim 23, including supporting a source overload condition, and
2   reducing source bias by randomizing the order of the sets of bids.